

The effect of numerical error on the reproducibility of molecular geometry optimizations

Christopher I. Williams · Miklos Feher

Received: 26 September 2007 / Accepted: 15 November 2007 / Published online: 6 December 2007
© Springer Science+Business Media B.V. 2007

Abstract Geometry optimization is one of the most often applied techniques in computational drug discovery. Although geometry optimization routines are generally deterministic, the minimization trajectories can be extremely sensitive to initial conditions, especially in case of larger systems such as proteins. Simple manipulations such as coordinate transformations (translations and rotations), file saving and retrieving, and hydrogen addition can introduce small variations (~ 0.001 Å) in the starting coordinates which can drastically affect the minimization trajectory. With large systems, optimized geometry differences of up to 1 Å RMSD and final energy differences of several kcal/mol can be observed when using many commercially available software packages. Differences in computer platforms can also lead to differences in minimization trajectories. Here we demonstrate how routine structure manipulations can introduce small variations in atomic coordinates, which upon geometry optimization, can give rise to unexpectedly large differences in optimized geometries and final energies. We also show how the same minimizations run on different computer platforms can also lead to different results. The implications of these findings on routine computational chemistry procedures are discussed.

Keywords Geometry optimization · Molecular mechanics · Numerical error

Introduction

Computational chemistry has now become an essential part of drug discovery. Although the approximate nature of computational models is well understood, it is usually implicitly assumed that the individual computational steps are both reproducible and ‘accurate’, at least within the level of approximations applied. Numerical errors are rarely discussed in computational chemistry, although it is well known that these errors are unavoidable when using finite precision computers for infinite aspects of mathematics [1, 2]. In many cases, numerical errors are small and have little effect on calculated properties. However, as the length and complexity of a calculation increases, numerical errors can accumulate to a point where both the calculated property and the reproducibility of the calculation can be drastically affected. As a result, many algorithms that are adequate for small systems can become unstable and fail on larger systems [3]. Computations that exhibit instabilities when faced with small input perturbations are known as being *sensitive to initial conditions*—(sometimes also referred to as ‘*chaotic*’) [4, 5]—each iterative cycle of such algorithms accumulate and magnify input errors to the point where the magnitude of those errors overwhelm the desired signal. Well-known examples of such systems include meteorological [4], seismological [6] and celestial trajectory simulations [7]. Numerical errors in timing calculations have also been blamed for foul-ups in missile guidance systems [8]. Many dissipative dynamic systems such as minimizations can exhibit *final state sensitivity* [9]—a condition characterized by fractal boundaries between attractive

C. I. Williams (✉)
Chemical Computing Group, Suite 910, 1010 Sherbrooke St. W.,
Montreal, QC, Canada H3A 2R7
e-mail: cw@chemcomp.com

M. Feher
Campbell Family Institute for Breast Cancer Research,
University Health Network, Toronto Medical Discovery Tower,
101 College Street, Suite 5-361, Toronto, ON, Canada M5G 1L7
e-mail: mfeher@uhnres.utoronto.ca

basins. Very small input perturbations near the basin boundaries of such systems will have drastic effects on the trajectory and final outcome of the calculation.

In computational chemistry, the chaotic nature of molecular dynamics (MD) trajectories and the resulting instabilities are well documented [10–13]. In MD simulations numerical instabilities often manifest themselves most strongly as a drift of the total energy of the system, leading to the effect of overheating, as well as the fact that even in 1–2 ps simulations very small perturbations of the atomic positions might lead to an exponential divergence of trajectories and to large differences in the resulting conformations. Other examples of numerical effects in computational chemistry include noisy regions produced by superposition of small exponent basis functions [14] and the tendency for self-consistent field (SCF) iterations to oscillate between solutions [15].

Recent years have seen the routine application of molecular mechanics geometry optimizations to increasingly large and more complex systems. The effects of numerical error are usually ignored in these calculations, and geometry optimizations are often assumed to be completely reproducible, as these computations typically contain no stochastic elements. However, as with any other complex calculation, propagation of numerical error through many iterative cycles of a minimization could cause the calculation to exhibit initial condition sensitivities, ultimately leading to divergent results when subjected to input perturbations or computer platform differences. Despite the well-known examples of numerical instability discussed above, and the fact that many computational chemistry applications could potentially exhibit initial state sensitivity, we currently know of no study that examines the effect of numerical errors on the reproducibility of molecular mechanics geometry optimizations. It is within this context that the authors began to study how small atomic coordinate perturbations and differences between computer platforms can affect the reproducibility such optimizations.

Numerical error effects in molecular mechanics geometry optimizations should be especially apparent when optimizing large molecules with highly complex potential energy surfaces such as proteins. As the number of particles (N) in the system increases, the spatial density of saddle points and local minima on the potential energy surface also increases and the MM potential energy surface (PES) becomes more complex and irregular than may be expected. It has been estimated [16] that, in the case of a Lennard-Jones gas, (i.e., van der Waals forces only) the number of possible minima (N_{MIN}) scales exponentially with the number of particles— $N_{\text{MIN}} \sim e^{\alpha N}$. Furthermore, the number of first order transition states (N_{TS1}) scales as $N_{\text{TS1}} \sim N e^{\alpha N}$ [17]. Here, α is a system-dependant parameter, estimates of which range from 0.02 to 13.

Even with low values of α , the sheer number of possible minima and higher order extrema suggests their spatial density must be quite high. As a result, even very small differences in starting coordinates can place structures on different sides of transition state cusps, and in the vicinity of different local minima. Although a protein is not Lennard-Jones gas, it is conceivable that the number of possible minima for a protein structure follows a similar exponential relationship with the number of atoms, leading to a large number of closely spaced extrema on the potential energy surface.

Small numerical differences in the starting atomic coordinates can be introduced from a number of simple and routine data manipulations, including the following sources:

- Coordinate errors due to inaccuracies in coordinate transformation (translation, rotation);
- Errors arising from finite field size (e.g. in PDB files) when saving/retrieving files;
- Errors arising from differences in hydrogen addition (e.g. when working from PDB files with no hydrogens).

In addition to input errors, digital calculations can be sensitive to the following sources of numerical error during run time:

- Errors arising from finite accuracy in molecular mechanics optimizations and finite final gradient sizes.
- The IEEE computing standard does not specify the order of all mathematical operations, so the compiler is free to determine this. Different operation order can cause very small changes to individual results.
- Simple changes to assembly level instructions will cause slightly different results. For instance, some computers have a multiply-add instruction that retain more accuracy than a multiply followed by and addition.
- When using multiple CPUs, numerical differences can also be introduced when making *data-dependent* branches (i.e., the flow of an algorithm is such that small data variations lead to different processors executing different pieces of code) or when race conditions between different calculation threads leads to changes in the sequence of operations [18].

Although it is quite difficult to separate out these errors, the aim of this work is to merely show that it is important to consider possible effects of these types of errors when working with large molecular systems.

In general it is difficult to know a priori whether or not a calculation will exhibit numerical sensitivity, but the behaviour can be probed empirically by testing the effects of input perturbations and computer platform differences

on the results of sample calculations [9]. Input error effects can be simulated by performing repeated runs on the same computer platform, each time using slightly perturbed versions of the input. Errors introduced by computer platform differences can be studied by running the exact same program and input on different computer systems. Calculations that do not exhibit sensitivity to initial conditions will produce identical or very similar results regardless of computer platform or input errors, while results produced from calculations that do exhibit sensitivity to initial conditions will vary substantially between computer platforms, and when subjected to small input perturbations. Molecular mechanics (MM) geometry optimizations are convenient for the empirical study of initial condition sensitivity because:

- (1) Optimization algorithms used by MM, such as Steepest Descent and Conjugate Gradient, are typically deterministic (no stochastic elements), so reproducibility is typically expected.
- (2) Data input errors can be introduced by small Cartesian coordinate perturbations or by saving to and retrieving from low precision file formats.
- (3) There exist programs that have been compiled on multiple platforms, so the effects of computer platform differences can be investigated.
- (4) Differences in geometry optimization trajectories accumulate because the coordinates in the $(n + 1)$ step of the algorithm (X_{n+1}) are computed using coordinates from the previous n step(s), X_n . In addition, the step size used in the optimization will also affect the results.
- (5) The complexity of the calculation can be increased by simply increasing the number of atoms in the system and/or the number of cycles in the simulation and/or the number of energy terms switched on in the potential.

- (6) Energy and gradient calculations involve spline approximations, trigonometric functions, $1/r^n$ terms and other ill-behaved mathematical forms that are sufficiently complex to exhibit sensitivity to numerical error.
- (7) Differences in results can be assessed quantitatively by comparing energies, gradients and coordinates, and qualitatively by visual inspection of structures.

With the above points in mind, we performed some simple tests to independently study the effects of input perturbations and computer platform differences on geometry optimizations. Input errors were explored by examining the effects of small atomic coordinate perturbations on the results of repeated optimizations using the same computer platform. Computer platform differences were explored by examining the results of optimizing the same molecular input using different computer platforms.

Experimental

Input errors from transformations and file I/O

To study input errors, perturbed input structures of six sample peptide systems (structures **1–6**, Table 1), varying in size from 44 to 3556 heavy atoms were considered. Structure **1** is a small peptide constructed using the MOE protein builder [19], minimized using the AMBER94 force field [20] in MOE, and saved to disk in PDB format. Structures **2–6** were generated by starting with the raw PDB files followed by deletion of bound waters and selected ligands. Heavy-atom only and hydrogen-added versions of each structures **2–6** were used as starting points for generating structures with small precision errors in the input coordinates. The MOE package was used to add the

Table 1 Structures considered in this study

Name	Sequence/PDB code	Number of amino acid residues	Heavy atom count	Protein class/name	Crystallographic resolution (Å)
1	ACE-ARG-PRO-HIS-PHE-NME	5	44	n/a	n/a
2	1LE8	51	414	MATA1/MATA1-alpha-2-3a heterodimer bound to DNA complex	2.30
3	1KF3	122	951	Atomic resolution structure of bovine pancreatic RNase A	1.05
4	1HDO	203	1544	Biliverdin-IX beta reductase	1.15
5	1KPI	289	2363	Mycolic acid cyclopropane synthase CMAA2 complexed with SAH and DDDMAB	2.65
6	1OJT	480	3556	Dihydropolipoamide dehydrogenase	2.75

Only amino acid chains were used; all ligands, oligo-peptides, oligo-nucleotides, solvent molecules, and non-covalently bound heterogroups were deleted before processing

hydrogens in all explicit hydrogen versions of the structures. For each of the six starting structures (heavy-atom only and hydrogen-added), ten coordinate-perturbed versions of the structure (**a–j**) were generated by subjecting each structure to a ‘random transformation’—a 10 Å translation pulse in a random direction followed by a random rotation. The random rotations employed quaternions rounded to the sixth decimal place, helping to introducing small but real differences in relative atomic coordinates. These new coordinates were then superposed back onto the original coordinates using the MOE pro_Superpose function, which determines the optimum superposition transformation between point sets by minimizing the mean square distance between the corresponding points. Because the superposition transforms were not *exactly* the reverse of the random transformations, the process yields structures with absolute coordinates *almost* identical to the starting structures, except for differences of ~ 0.0001 Å in the atomic coordinates. Further loss in coordinate precision was introduced by writing the new coordinates to disk in the PDB format, which supports only a limited precision for the coordinates. The new PDB structure files thus produced are identical to the original structures except for minute differences (~ 0.001 Å) in the atomic coordinates. The ten (10) perturbed versions of structures **1–6** were used for subsequent numerical sensitivity tests. In the heavy-atom only versions of structures **2–6**, adding hydrogens to fill valence was performed after the PDB files were read into the molecular modeling packages—this was meant to mimic a typical workflow scenario, and adds additional starting coordinate differences to these structures that arise from hydrogen placement. Details on structures **2–6** are given in Table 1.

Effect of input errors on MM optimizations: structure (**1**)

Each of the 10 perturbed versions of structure **1** was read into Hyperchem [21], ChemX [22], MOE and Cerius2 [23] packages, and subjected to a molecular mechanics minimization down to a root mean square (RMS) gradient of 10^{-5} kcal/mol Å². (The packages MacroModel and Discovery Studio were not involved in these tests because hydrogen-filled PDB structures would have required some manual atom typing after reading in the structures, and this could have introduced additional errors into the results.) Geometry optimizations were performed using the conjugate gradient method, whereas in MOE the default minimization protocol, that employs a cascade of minimization routines starting with steepest descent, followed by conjugate gradient and truncated Newton, was applied. The

following MM force fields were applied: the MMFF94s [24] force field in MOE with the default settings, MM+ in HyperChem, ChemX force field in ChemX and the Dreiding [25] force field in Cerius2.

Effect of input errors on MM optimizations-heavy atom only structures (**2–6**)

Geometry optimizations on heavy-atom only structures **2–6** were performed using the Cerius2, MOE, Discovery Studio (DS) [26] and MacroModel [27] software packages. The Cerius2 minimizations were performed using stringent ‘high convergence’ settings: (RMS force on atoms 10^{-3} kcal/molÅ, maximum force on atoms 0.005 kcal/molÅ, overall energy difference between steps 10^{-4} kcal/mol, overall rms displacement 10^{-5} Å, maximum displacement 5×10^{-5} Å). It must be noted that the results didn’t substantially differ from those obtained using the less stringent ‘normal precision’ settings. The MOE optimizations were performed with the default optimizations settings except for the convergence criterion, which was lowered to RMS gradient of 10^{-5} kcal/mol Å². The Discovery Studio minimizations were performed using the CHARMM force field [28] with the ‘Adopted Basis NR’ algorithm to an RMS Gradient of 10^{-8} . No implicit solvent model was used and the dielectric constant was set to 1, otherwise default conditions were applied. In MacroModel, geometry optimizations were carried out using the OPLS2005 force field [29] using the Powell-Reeves conjugate gradient (PRCG) method with default parameters and constant dielectrics and no solvent. The calculations were terminated when the gradient was below 10^{-4} .

Effect of input errors on MM optimizations-hydrogen added structures (**2–6**)

Geometry optimizations on hydrogen-added structures **2–6** were performed using the MOE, Discovery Studio (DS) [26] and MacroModel [27] software packages. The minimization criteria and forcefield setting were the same as those used for the heavy-atom only structures.

Effects of platform differences on MM optimizations

The reproducibility of MM calculations across computer platforms was studied with MOE and Discovery Studio, because both programs have been compiled on different computer platforms. Furthermore, the MOE software uses the same underlying C code for the potential calculation on all supported platforms. With these tests, differences in

optimizations on different platforms can be attributed strictly to software and architecture differences between the computer platforms. The ten perturbed versions of structures **1** and **3** were minimized in MOE on a three different computer platforms; a 1 GHz Pentium4 Intel clone with 256 Mb of RAM running Windows 2000 service pack 4 operating system (henceforth referred to as the “MOE-Windows” system) a Silicon Graphics Octane2 with a 400 MHz IP30 processor (CPU: MIPS R12000 Processor Chip revision 3.5; FPU MIPS R12010 Floating point chip revision 0.0) and 512 Mb of RAM running IRIX 6.5 (henceforth referred to as the “MOE-SGI” system) and an IBM PowerPC 9113–550 with a Quad processor and 4 Gb of RAM running AIX 5.2 (henceforth referred to as the “MOE-IBM AIX” computer). The ten perturbed versions of structure **3** were minimized in Discovery Studio on an HPxw8200 with 2 Intel Pentium4 cpu’s at 3.2 GHz with 3Gb RAM running under Windows XP Service Pack 2 (henceforth referred to as the “DS-Windows” system) and on a Sun Fire V40z server with 8 dual-core ADM Opteron processors running under RedHat Enterprise Linux, release 4 (henceforth referred to as the “DS-Linux” system).

Results and discussion

The initial and final molecular mechanics energies for ten coordinate-perturbed versions of structure **1** are listed in Table 2. Table 3(a, b) provide summaries of the corresponding results for heavy-atom only and hydrogen-added structures **2–6**. The heavy-atom only starting structures were minimized using the MOE, Cerius2, Discovery Studio

and MacroModel programs. The hydrogen added structures were minimized with all of the packages except Cerius2.

Input errors from transformations and file I/O

The average RMSDs arising from translation/rotation operations before writing to disk are small but real ($<10^{-5}$ Å), and reflect truncation and rounding errors incurred during the transform calculations. The average coordinate RMSD difference after reading back from the PDB disk files are substantially larger (~ 0.001 Å) and are comparable in magnitude to the allowed precision in the PDB format. Either way, the input errors introduced by these operations seem trivial, and they are much smaller than the maximum precision that can be hoped for from experimental coordinates. The RMSD errors show little size effect and appear to maintain the same order of magnitude when going from 44 to 3556 heavy atoms (results not shown).

Effect of input errors on MM optimizations

Since the MM optimization routines employed here contain no stochastic elements, repeated minimization runs of an identical starting structure performed on one computer platform with one CPU are expected follow the exact same optimization trajectory to the exact same nearest local minimum. We tested this hypothesis by performing the optimization of structure **1** ten times in both the Hyperchem and MOE programs, and structure **2** ten times in MOE alone; as expected, the energies and gradients at each optimization step were identical to all decimal points of

Table 2 MM optimizations of structures **1a–1j**^a

Structure	Hyperchem		ChemX		MOE-Windows		Cerius2	
	E initial	E final	E initial	E final	E initial	E final	E initial	E final
1a	141.9467	38.6570	551.8490	84.2568	67.55239	17.66567	265.8100	58.0338
1b	141.9468	38.6430	551.8490	86.2627	67.50687	17.66567	265.8330	58.0566
1c	141.9468	38.6075	551.8490	85.4776	67.66428	17.66567	265.8410	58.0648
1d	141.9466	38.6666	551.8490	85.4261	67.49571	17.66567	265.8100	58.0338
1e	141.9468	38.6192	551.8490	85.3341	67.50810	17.66567	265.8100	58.0338
1f	141.9467	38.6344	551.8490	83.9043	67.60683	17.66567	265.8330	58.0648
1g	141.9467	38.6471	551.8490	82.3829	67.58240	17.55958	265.8410	58.0338
1h	141.9468	38.5278	551.8490	86.3013	67.58289	17.55958	265.8100	58.0566
1i	141.9468	38.5479	551.8490	86.2702	67.71867	17.66567	265.8330	58.0338
1j	141.9467	38.4266	551.8490	86.2591	67.70234	17.66567	265.8100	58.0338
Range	0.0000	0.2400	0.0000	3.9184	0.22296	0.10609	0.0310	0.0310

^a Structures **1a–1j** were generated by coordinate transformations from an initial seed structure. All energies in kcal/mol. See text for further details

Table 3 MM optimizations of structures **2–6**: (a) from heavy-atom only PDB structures; Hydrogens added in the respective packages^a; (b) from PDB files with hydrogens previously added^a

Software package	Structure	E initial (kcal/mol)			RMSD (Å)	E final (kcal/mol)			RMSD (Å)
		Ave	SD	Range		Ave	SD	Range	
(a) From heavy-atom only PDB structures; Hydrogens added in the respective packages									
C2	2	2691.863	0.773	2.240	0.001	428.758	17.591	56.820	0.183
	3	2416.676	1.791	4.680	0.001	-17.075	11.332	33.440	0.210
	4	9761.490	26.518	86.940	0.001	100.397	22.453	66.140	0.310
	5	9661.692	30.887	91.010	0.001	-285.579	15.587	49.990	0.310
	6	250735.800	1262.776	3818.000	0.001	797.257	32.376	100.290	0.429
MOE	2	1135.834	0.548	1.710	0.001	-30.350	1.862	8.125	0.223
	3	2373.917	0.334	1.192	0.001	-63.117	10.341	32.305	0.626
	4	4635.215	1.905	5.965	0.001	-278.154	31.389	94.167	0.808
	5	5866.655	1.074	3.806	0.001	-820.457	55.491	155.148	0.793
	6	16402.879	4.895	16.697	0.001	-222.845	78.416	278.178	1.181
Discovery studio	2	-1110.384	1.137	3.388	0.001	-3064.303	41.857	116.515	0.699
	3	-3955.009	0.494	1.593	0.001	-8074.333	64.436	218.931	0.705
	4	-5572.740	1.671	5.877	0.001	-12831.703	68.395	263.422	0.764
	5	-624.520	83.307	308.997	0.001	-20583.117	106.800	333.066	0.944
	6	67577.328	1577.258	429.578	0.001	-30971.718	117.650	391.527	1.115
Macro Model	2	209.388	0.733	2.368	0.001	-1429.142	5.060	22.001	0.965
	3	-2196.953	0.860	2.868	0.001	-4687.989	27.678	77.702	0.833
	4	-1025.237	1.258	3.842	0.001	-6390.147	30.470	107.845	1.879
	5	-6428.480	0.743	2.326	0.001	-14361.578	30.371	109.706	1.810
	6	3948.029	10.716	37.386	0.001	-20924.318	43.681	147.484	1.309
(b) From PDB files with hydrogens previously added									
MOE	2	1135.832	0.549	1.720	0.001	-30.718	1.585	5.362	0.129
	3	2373.915	0.334	1.190	0.001	-56.690	7.460	28.591	0.515
	4	4635.216	1.905	5.970	0.001	-274.723	24.008	87.965	0.838
	5	5866.655	1.075	3.810	0.001	-828.957	41.549	142.071	0.796
	6	16420.881	4.888	16.701	0.001	-215.114	80.899	260.773	1.130
Discovery studio	2	-1407.238	0.255	0.825	0.001	-3100.460	34.010	100.311	0.817
	3	-4163.606	0.325	0.963	0.001	-8087.243	75.412	245.436	0.890
	4	-6253.759	0.403	1.584	0.001	-12831.999	63.222	213.586	0.492
	5	-7580.998	2.290	8.457	0.001	-20582.447	145.452	398.090	0.837
	6	23548.428	33.900	118.410	0.001	-30909.203	200.409	807.313	0.953
Macro Model	2	-948.666	0.186	0.571	0.001	-2114.814	7.143	23.477	0.663
	3	-3444.996	0.483	1.510	0.001	-5324.197	0.154	0.497	0.082
	4	-5346.976	0.424	1.621	0.001	-8733.273	13.728	46.645	0.840
	5	-8175.005	1.366	3.775	0.001	-13682.412	26.693	74.476	0.668
	6	332.268	3.151	9.047	0.001	-20782.226	28.815	108.389	1.475

^a The following statistical parameters were calculated: Ave—mean of the calculated energies; SD—their standard deviation. RMSD is the root-mean-square atomic distance calculated for the heavy atoms of the protein structures after superposition using their α -carbons. See text for further details

precision in each of the repeated runs. Thus, the optimizations all followed exactly the same trajectory to exactly the same local minimum. This procedure was repeated for all structures **3–6** with the same results; this suggests that in the confines of one computer platform and 1 CPU,

minimizations on identical starting structures are completely reproducible.

Differences between the MM trajectories begin to appear as small perturbations are introduced into the starting structure. In Table 2, the initial energies of

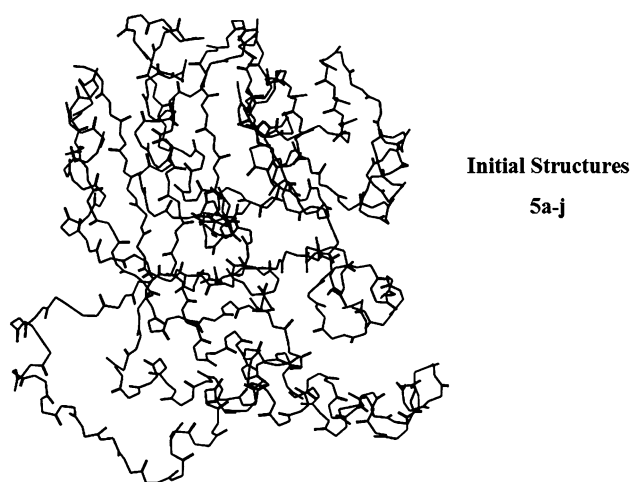


Fig. 1 Overlay of 10 randomly perturbed protein structures before minimization. (The structure is 1KPI in the PDB, shown as structure **5** in Tables **1** and **3a–b**). Only the backbone atoms are shown. See text for details

structures **1a–1j** show that the small coordinate errors introduced by the transformations and I/O operations can have small effects on the initial MM energies (E_{initial}). The variation in sensitivity to numerical error exhibited by the single-point E_{initial} energies of different forcefields depends primarily on the gradient of the forcefield at that point in phase space. Calculations at potential energy surface (PES) points with large gradients will be significantly affected by small coordinate perturbations, while small coordinate perturbations will have little effect in regions of the PES where the gradient is small.

Subsequent minimizations of structures **1a–1j** show an interesting result; the differences in the minimized energies (E_{final}) can be greater than the differences in the starting energies (E_{initial}). Also, the individual software packages behave quite differently. ChemX and Hyperchem show identical and near-identical initial MM energies despite the small coordinate perturbations, but subsequent MM optimizations with these packages leads to optimized structures with a range of final energies larger than the range of initial energies. In contrast, MOE shows larger differences in the initial (unoptimized) energy values than either ChemX or Hyperchem, but somewhat smaller energy differences after optimization. Using the Cerius2 package, the range of energies is the same before and after the MM optimization. A possible explanation for the behaviour of C2 might be that the results reflect the simplicity of the Dreiding potential energy surface compared to the other more complex force fields studied here. In general, the sensitivity of a forcefield to numerical error is expected to depend on its complexity, and more precisely, the roughness of its potential energy surface and the density of local minima it produces in phase space. Forcefields with smooth potential energy surfaces and relatively sparse local minima will

show less numerical sensitivity than forcefields with rough potential energy surfaces and densely packed local minima. Factors that could also affect the numerical sensitivity of a minimization include the numerical sensitivity of the optimization routine(s) employed and possibly even the compiler options and math libraries used to compile the program. However, the effect of these components is more speculative, and the underlying sensitivity to numerical error probably arises mainly from a rough potential energy surface and a high density of local minima.

The ChemX and Hyperchem programs both produce a different final energy for each perturbed version of structure **1**, suggesting that each input structure optimizes to a unique minimum which is similar, but not identical to, the minima found by the other starting structures. In contrast, the MOE and C2 optimizations produce final structures that converge to either two distinct minima when using MOE ($E_{\text{final}} = 17.6657$ or 17.5596 kcal/mol) or three distinct minima when using C2 ($E_{\text{final}} = 58.0338$ or 58.0566 or 58.0648 kcal/mol). One could argue that deeper minimization of these structures to even smaller gradients may result in convergence towards a single energy and structure. However, this is not the case; even if the final convergence criteria are made 2–3 orders of magnitude stricter, the range in the final energies hardly changes. The interesting observation here is the unexpected sensitivity of the minimizations to small starting position perturbations, ultimately causing the process to end up in different local minima.

The peptide structure **1** is a relatively small system, and the energy differences in the final structures are inconsequential when compared to protein-ligand binding energies, which typically range from ~ 2 – 15 kcal/mol [30]. However, as the molecular systems get larger, the variations in the optimized energies and geometries become substantial. The results for the proteins **2–6** are summarized in Table **3a** (heavy-atom only source structures) and Table **3b** (hydrogen-added source structures). The tables containing the average initial and final energies, as well as their corresponding average errors and ranges. These tables illustrate that for larger systems the range of obtained energies generally increases with system size, and can become quite substantial. In the case of structure **6**, the final energies calculated from different starting coordinates can differ by up to 100 kcal/mol. Also, in the case of structures **2–6**, it is rare that within the ten examples a given final energy is repeated. This behaviour is quite different from the ‘expected’ reproducibility. It must be noted that we observed similar effects when optimizing molecular geometries using semi-empirical quantum chemistry at the AM1 and PM3 levels (results not shown).

Qualitatively, the RMSDs between the non-optimized structures are so small as to not be visible in line mode

Fig. 2 Overlay of 10 randomly perturbed protein structures after minimization using low final gradient settings in four different computer programs. (The structure is 1KPI in the PDB, shown as structure **5** in Tables 1 and 3a–b). Only the backbone atoms are shown. See text for details

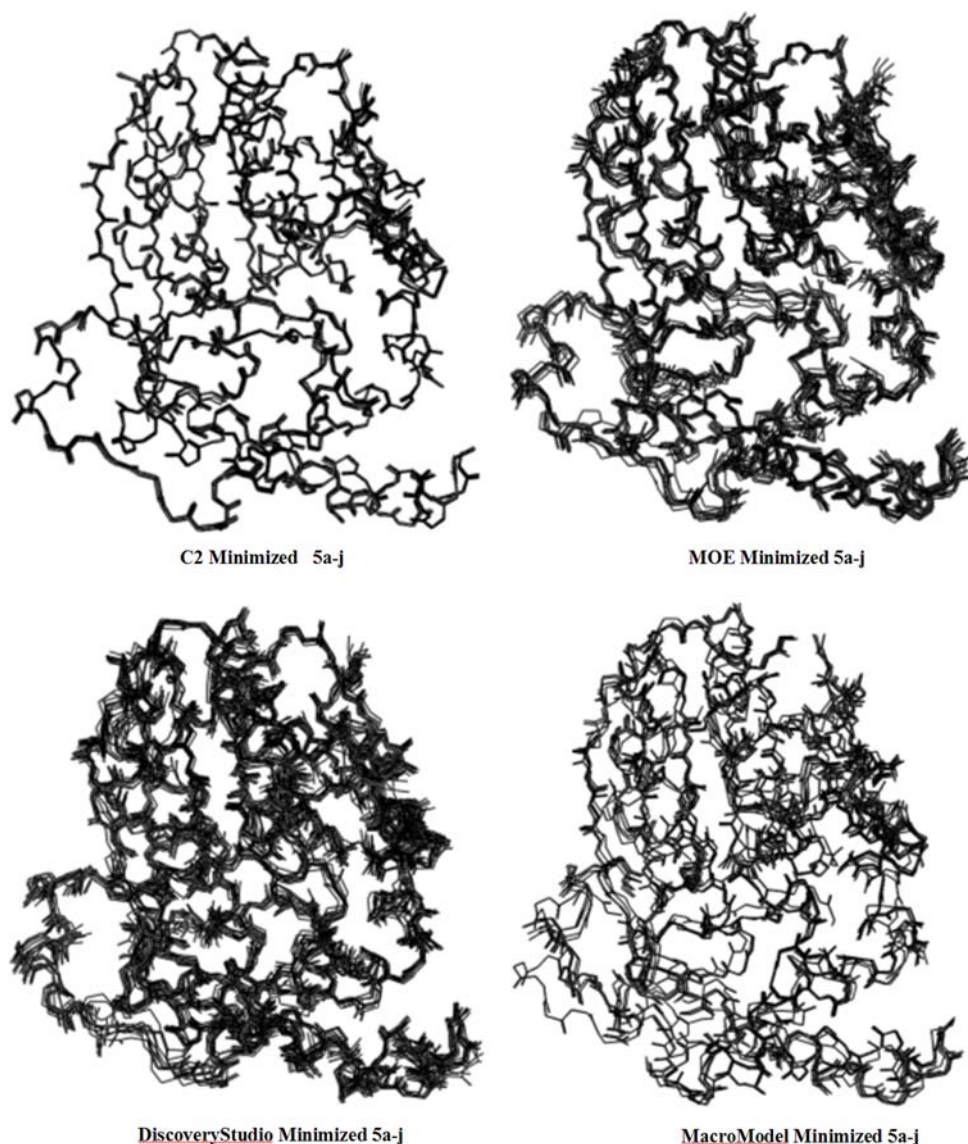


Table 4 Breakdown of the average pairwise RMSD for structure **5** (mycolic acid cyclopropane synthase 1KPI)^a

Software package	Average pairwise RMSD (Å)				
	Backbone atoms	Sidechain atoms	Helical regions	Sheet regions	Loops and disordered regions
C2	0.2454	0.2988	0.2476	0.2327	0.3113
MOE	0.4797	0.6031	0.5122	0.3338	0.6055
Discovery studio	0.6527	0.8309	0.7505	0.4900	0.7135
Macro Model	0.9103	1.2398	1.0966	0.7932	1.1192

^a The average RMSD values were calculated after all-atom based superposition of the optimized structures. The RMSD calculations considered all of the heavy atoms in the each of the structural subsets

rendering of the overlaid structures, as shown for structure **5** in Fig. 1. A similar superposition of the Cerius2, MOE, Discovery Studio and MacroModel optimized versions of structure **5** (Fig. 2) reveals obvious differences between the optimized structures, with significant structural variation in some regions in the protein.

Detailed RMSD analysis of the optimized structures **2–6** shows that on average the sidechains show greater deviation in the optimized structures than the backbone. In Table 4 the average pairwise RMSD between the optimized structure **5** geometries is broken down into contributions from the sidechain atoms, backbone atoms,

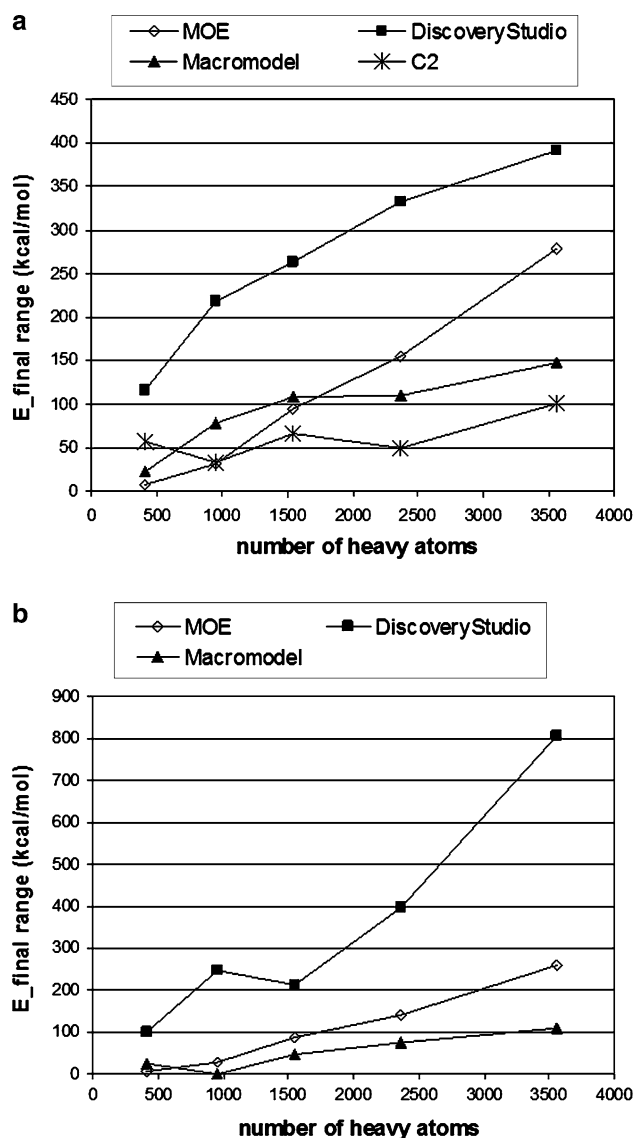


Fig. 3 Plot of the final energy range after minimization versus the number of heavy atoms for different protein structures. (a) For each of the proteins ten perturbed versions were generated using coordinate transformations. The molecules were read into the respective packages from PDB files, hydrogens were added and the structures minimized to a low gradient. (b) For each of the proteins ten perturbed versions were generated using coordinate transformations and then hydrogens were added. The molecules were read into the respective packages from PDB files and minimized to a low gradient. See text for further details

helices, sheets and loops/disordered regions. The results show that in general the sidechain, loops and disordered regions show the greatest deviations between the optimized structures. This is expected because these atoms are in more disordered and peripheral regions of the protein, where the minima on the potential energy surface are more broad and shallow.

The variation in final energies and geometries of structures 2–6 is somewhat unexpected, considering that the

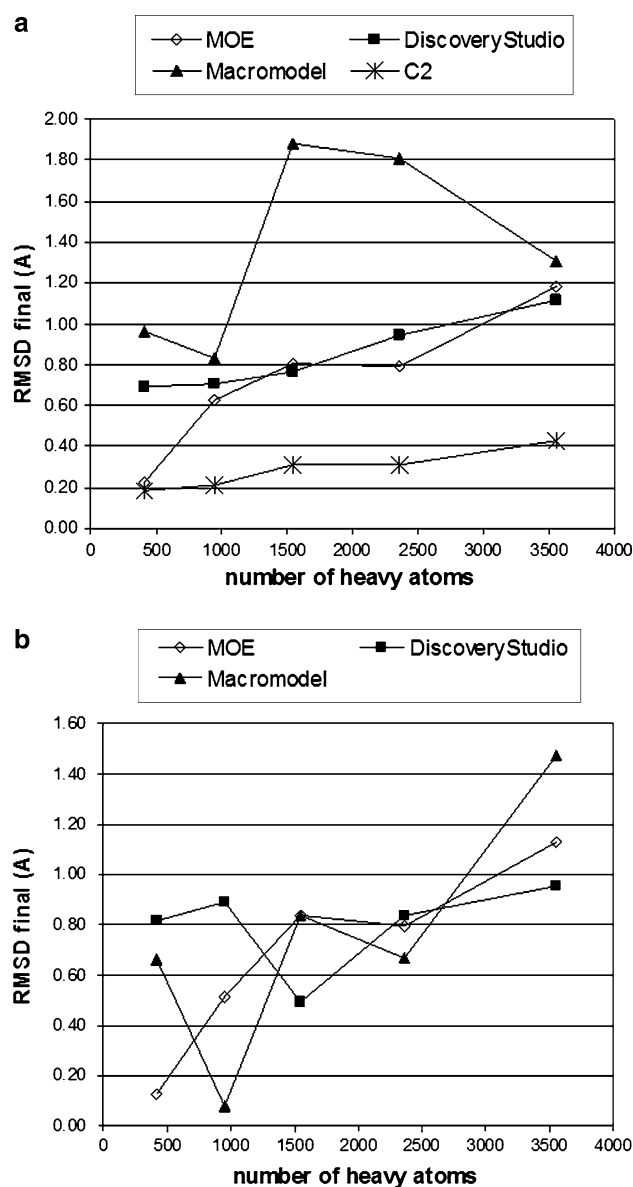


Fig. 4 Plot of the average pairwise RMSD (Å) after minimization versus the number of heavy atoms for different protein structures. (a) For each of the proteins ten perturbed versions were generated using coordinate transformations. The molecules were read into the respective packages from PDB files, hydrogens were added and the structures minimized to a low gradient. (b) For each of the proteins ten perturbed versions were generated using coordinate transformations and then hydrogens were added. The molecules were read into the respective packages from PDB files and minimized to a low gradient. See text for further details

coordinate errors introduced into the starting structures are all quite small. Although these input errors were purposefully created, it is important to recognize that these input differences are realistic, and can be inadvertently introduced during commonplace computational chemistry structural manipulations and structure output/retrieval from physical disk. It is also important to recognize that the

Table 5 MOE minimizations of structure **1** on different computer platforms^a

Run	MOE-Windows		MOE-SGI		MOE-IBM AIX	
	E initial	E final	E initial	E final	E initial	E final
1a	67.55239	17.66567	67.55240	17.66567	67.55240	17.66568
1b	67.50687	17.66567	67.50687	17.66567	67.50687	17.66568
1c	67.66428	17.66567	67.66428	17.66567	67.66428	17.66568
1d	67.49571	17.66567	67.49571	17.66567	67.49571	17.66568
1e	67.50810	17.66567	67.50811	17.66567	67.50811	17.66568
1f	67.60683	17.66567	67.60684	17.66567	67.60684	17.66568
1g	67.58240	17.55958	67.58241	17.55958	67.58241	17.55959
1h	67.58289	17.55958	67.58290	17.55958	67.58290	17.55959
1i	67.71867	17.66567	67.71867	17.66567	67.71867	17.66568
1j	67.70234	17.66567	67.70235	17.66567	67.70235	17.66568
Energy range (kcal/mol)	0.22296	0.10609	0.22296	0.10609	0.22296	0.10609
Rms distance (Å)	0.00010	0.00120	0.00010	0.00010	0.00010	0.00090

^a Structures **1a–1j** were generated by coordinate transformations from an initial seed structure. All energies in kcal/mol. See text for further details

degree of variation in both final energy and RMSD between optimized structures increases with system size. In Fig. 3(a, b) plots of the range of final energies as a function of increasing heavy atom count show that in general all the packages show an increase in final energy range with increasing heavy atom count, with both the heavy-atom and hydrogen-added source structures. The plots in Fig. 4(a, b) shows that the final pairwise RMSD also increases with heavy atom count, although the rate of increase is not completely uniform.

The initial energy differences between the heavy-atom only and the hydrogen-added source structures (Table 3a, b) reflect differences in hydrogen placement between the software packages. Since MOE was used to add hydrogens to all the hydrogen-added structures, very little difference is seen between the initial MOE energies of the heavy-atom only and hydrogen-added structures. In contrast, the difference in initial energy of the heavy-atom only and hydrogen-added structures is larger for the other packages, reflecting small differences between the hydrogen addition routine in MOE and the hydrogen addition routine in each respective package. These hydrogen placement differences could produce additional energy variations during optimization, but in most cases there is little difference between the variation in final energies of the heavy-atom only and hydrogen-added source structures. The only exception to this trend is the Discovery Studio optimization of structure **6**, where the variation in final energy of the optimized hydrogen-added source structures is four-fold the energy variation of the heavy-atom only source structures. A possible explanation for this behavior is that the hydrogen positions produced by MOE are sufficiently different from those produced by Discovery Studio that the

hydrogen-added starting structures are in a much higher gradient region of the potential energy surface than the heavy-atom only starting structures where the hydrogens were added in Discovery Studio. Larger gradients in the beginning of the calculation would increase the sensitivity of the optimization to chaotic effects.

Effects of platform differences on MM optimizations

The initial and final energies obtained on three platforms for structure **1** using MOE are listed in Table 5. For a given perturbed structure (e.g. **1b**), the initial energy using MOE varies less than 0.001 kcal/mol across platforms—much smaller variation than was introduced by the coordinate perturbations. This lack of variation in initial energy across computer platforms is to be expected, because the exact same input file is used in each case, and only one single-point energy calculation is performed to compute the initial energy. Thus, energy differences at this point are solely the result of differences in mathematical function evaluation between the computer platforms. Upon minimization of structure **1** with MOE, two final energy states were produced on all three platforms, i.e. the sensitivity on the starting geometry was reproduced, but no major platform dependence was seen. Furthermore, the two final energies are identical to those produced by the coordinate perturbed structures on a single platform, adding additional support to the observation that these structures minimize to two distinct minima.

The initial and final energies of structure **3** obtained on three platforms using MOE and two platforms using Discovery Studio are listed in Tables 6(a, b). As with structure **1**,

Table 6 (a) MOE minimizations of structure 3 (bovine pancreatic RNase A 1KF3) on different computer platforms; (b) Discovery studio minimizations of structure 3 (bovine pancreatic RNase A 1KF3) on two computer platforms^a

Run	MOE-Windows		MOE-SGI		MOE-IBM AIX	
	E initial	E final	E initial	E final	E initial	E final
(a) MOE minimizations of structure 3 (bovine pancreatic RNase A 1KF3) on different computer platforms						
3a	2373.24268	-51.40294	2373.24268	-50.91029	2373.24285	-48.79851
3b	2374.20581	-50.69243	2374.20581	-50.69226	2374.20599	-50.69223
3c	2373.75903	-51.40294	2373.75903	-48.79851	2373.75929	-58.45880
3d	2374.43457	-58.85843	2374.43481	-50.65509	2374.43487	-66.54001
3e	2374.20044	-57.52506	2374.20068	-58.22276	2374.20079	-57.52482
3f	2373.67310	-75.18790	2373.67334	-73.99396	2373.67340	-73.61921
3g	2373.65479	-60.44793	2373.65479	-61.98717	2373.65497	-61.96611
3h	2373.84277	-46.59679	2373.84277	-52.57420	2373.84292	-48.47718
3i	2373.97900	-57.22924	2373.97925	-55.13151	2373.97930	-52.76701
3j	2374.17212	-57.55516	2374.17212	-63.64044	2374.17230	-67.60210
Energy range (kcal/mol)	1.19189	28.59111	1.19213	25.19545	1.19202	25.14202
rms Distance (Å)	0.00120	0.51450	0.00120	0.56730	0.00120	0.54300
Run	DS-Windows		DS-Linux			
	E initial	E final	E initial	E final		
(b) Discovery studio minimizations of structure 3 (bovine pancreatic RNase A 1KF3) on two computer platforms						
3a	-3955.52100	-8009.59277	-3955.52100	-8096.28418		
3b	-3954.41772	-8149.74805	-3954.41772	-8153.10303		
3c	-3954.08301	-8109.54834	-3954.08301	-8184.93018		
3d	-3954.70020	-7930.81738	-3954.70020	-8048.11475		
3e	-3955.67627	-8052.40137	-3955.67627	-8081.83496		
3f	-3955.40625	-8121.78076	-3955.40625	-8180.61963		
3g	-3955.44336	-8037.54346	-3955.44336	-8087.85938		
3h	-3955.11841	-8128.22656	-3955.11841	-7993.68115		
3i	-3954.75366	-8127.34033	-3954.75366	-8164.90869		
3j	-3954.96460	-8076.32764	-3954.96460	-8104.40137		
Energy range (kcal/mol)	1.59330	218.93070	1.59326	191.24903		
rms Distance (Å)	0.001	0.705	0.001	0.775		

^a Structures **3a–3j** were generated by coordinate transformations from an initial seed structure. All energies in kcal/mol. See text for further details

the initial MOE energies of a given perturbed version of structure **3** varies less than 0.001 kcal/mol across platforms. Using Discovery Studio, the initial energy of a given perturbed structure is identical across platforms to the five decimal places reported in the Discovery Studio log file. However, in contrast with structure **1**, the final energy of a given perturbed version of structure **3** varies substantially between platforms. With MOE, the final energy of a given starting structure can vary by 10 kcal/mol across the computer platforms; with Discovery Studio, the cross-platform variation in final energies can be as high as 100 kcal/mol. It should be noted that these energies are much larger than typical protein-ligand binding energies. The results for structures **1** and **3** suggest that the final

results from geometry optimizations will be increasingly platform dependent as the size of the molecular system increases.

Conclusions

The behaviour of geometry optimization calculations when faced with perturbed input structures, different computer platforms and different programs suggests that sensitivity to initial conditions might be a common problem in molecular mechanics minimizations. With large systems, the errors produced as a result of this sensitivity can be of sufficient magnitude to affect the qualitative and

quantitative conclusions drawn from the results. The sensitivity of these geometry optimizations calls into question what one really means by the “nearest local minimum” in an MM optimization. In this regard, the situation is somewhat similar to molecular dynamics, where “it has been a common frustrating experience that when a computer or compiler has been slightly changed, trajectories of MD cannot be reproduced” [31]. It now appears that simple MM optimizations also have similar properties. The root cause of this effect, non-linear interactions inherent in molecular mechanics force fields, has been identified and discussed in the context of molecular dynamics simulations [32]. We have now shown that in practice, even a straightforward energy optimization under certain circumstances cannot be used in a deterministic manner (even though the process itself is fully deterministic), due to its high sensitivity to input precision. Geometry optimizations are often viewed as a marble rolling down a bowl—a picture arising from the harmonic representation of the diatomic potential energy curve. In this scenario, small differences in starting position result in at most small differences in final energies, and these differences can be made infinitesimally small by improving the precision of the calculation. However, as system size increases, minimizations become more akin to a pebble falling down a rocky hillside; small perturbations in initial positions can lead to large differences in the path taken and the final resting place on the valley floor. In this case, no improvement in the precision of the calculations will substantially reduce the large differences in final energy.

Although it is important to recognize input sensitivity as a potential source of error and ambiguity, it is difficult to propose a practical solution. Perhaps the most akin to the issues described in this work are studies of the chaotic nature of molecular dynamics trajectories in proteins [32]. It was shown that in molecular dynamics simulations very small perturbations of the atomic positions (10^{-3} – 10^{-9} Å) led to an exponential divergence of trajectories and to conformations differing by as much as 1 Å RMSD within an elapsed time of 1–2 ps [31]. The effect of these issues on protein folding calculations has also been demonstrated [33]. It was somewhat pessimistically concluded that “individual MD trajectories of folding are too sensitive to small perturbations to have significant predictive quality” [31]. Luckily, the situation appears to be a lot less serious in geometry optimizations. However, to deal with numerical instability, one could ideally map the potential surface in the interesting region. As this is impractical for large molecules such as proteins, it might be a possibility to purposely generate a small number of closely lying starting points (e.g. by coordinate perturbation) and perform the minimization for all of these, selecting the most appropriate solution (e.g. the lowest energy one or perhaps some

kind of an average). When failing to use multiple starting points, however, it is very important to bear in mind that these errors might be quite significant. This will especially be the case when calculating small differences in large energies (e.g. when determining protein-ligand binding energies from the energy difference between the complex and the free ligand and protein), or when the initial gradient is substantial (e.g. when minimizing a docked structure in the field of the receptor, or especially when a ligand is placed inside a receptor with no prior bound ligand). In such cases, the errors might be of the same order of magnitude as the calculated quantity, and researchers need to be aware that performing complex calculations under slightly different conditions or different platforms can lead to significantly different results.

Acknowledgements The contribution of Eugen Derety in the early stages of this work is acknowledged. Paul Labute, Martin Santavy and Jocelyn Demers of Chemical Computing Group are gratefully acknowledged for insightful discussions about the subject. James Bugden is also thanked for useful comments on numerical stability and compiler differences.

References

1. Scheid F (1968) Theory and problems of numerical analysis. McGraw-Hill Book Company, New York
2. Hayes B (2003) A lucid interval. *Am Scientist* 91:484–488
3. Demmel J (1992) Trading off parallelism and numerical stability. Computer Science Division Tech Report UCB//CSD-92-702, University of California, Berkeley
4. Lorenz E (1996) The essence of Chaos. University of Washington Press, Seattle. ISBN: 0295975148
5. Gleick J (1987) Chaos: making a new Science. Penguin Books, New York. ISBN: 0140092501
6. Sornette D (2002) Nature debates: earthquakes 1999. Internet References. Retrieved from <http://www.nature.com/nature/debates/earthquake/>, accessed on 9/19/2002
7. Groison D (2002) Est-il vrai que les ordinateurs font des erreurs de calcul? *Science Vie* 1022:130
8. Skeel R. (1992) Roundoff error and the patriot missile. *SIAM News* 25:11
9. McDonald SW, Gregbogi C, Ott E, Yorke JA (1985) Fractal basin boundaries. *Physica D* 17:125–153
10. Barth E, Schlick T (1998) Extrapolation versus impulse in multiple-time stepping schemes. II. Linear analysis and applications to Newtonian and Langevin dynamics. *J Chem Phys* 109:1633–1642
11. Biesiadecki JJ, Skeel RD (1993) Dangers of multiple time step methods. *J Comput Phys* 109:318–328
12. Bishop TC, Skeel RD, Schulten K (1997) Difficulties with multiple time stepping and fast multipole algorithm in molecular dynamics. *J Comput Chem* 18:1785–1791
13. Sandu I A, Schlick T (1999) Masking resonance artifacts in force-splitting methods for biomolecular simulations by extrapolative Langevin dynamics. *J Comput Phys* 151:74–113
14. Hatano Y, Yamamoto S, Tatewaki H (2005) Characterization of molecular orbitals by counting nodal regions. *J Comput Chem* 26:325–333
15. Young DC (2001) Computational chemistry: a practical guide for applying techniques to real-world problems. Wiley, New York

16. Stillinger FH (1999) Exponential multiplicity of inherent structures. *Phys Rev E* 59:48–51
17. Doye JPK, Wales DJ (2002) Saddle points and dynamics of Lennard-Jones clusters, solids and supercooled liquids. *J Chem Phys* 116:3777–3788
18. Blackford LS, Cleary A, Petitet A, Whaley RC, Demmel J, Dhillon I, Ren H, Stanley K, Dongarra J (1997) Practical experience in the numerical dangers of heterogeneous computing. *ACM Trans Math Software* 23:133–147
19. MOE (2004) Molecular Operating Environment, version 2004.03; Chemical Computing Group Inc., Montreal
20. Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM, Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA (1995) A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J Am Chem Soc* 117:5179–5197
21. HyperChem, version 6.0, HyperCube Inc., Gainesville, FL, 32601, USA
22. ChemX, version 2000.1, Oxford Molecular Ltd., currently Accelrys Inc., San Diego, CA, 92121, USA
23. Cerius2, version 4.9, 2004, Accelrys Inc., San Diego, CA, 92121, USA
24. Halgren TA (1996) Basis, form, scope, parametrization, and performance of MMFF94. *J Comput Chem* 17:490–519
25. Mayo SL, Olafson BD, Goddard WA (1990) DREIDING: a generic force field for molecular simulations. *J Phys Chem* 94:8897–8909
26. Discovery Studio, version 1.7, 2007, Accelrys Inc., San Diego, CA, USA
27. MacroModel 9.1, Schrodinger Inc., Portland, OR, 97204, USA
28. Brooks BR, Brucoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M (1983) A program for macromolecular energy, minimization, and dynamics calculations. *J Comput Chem* 4:187–217
29. Jorgensen WL, Maxwell DS, Tirado-Rives J (1996) Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J Am Chem Soc* 118:11225–11236
30. Kuntz ID, Chen K, Sharp KA, Kollman PA (1999) The maximal affinity of ligands. *Proc Natl Acad Sci USA* 96:9997–10002
31. Braxenthaler M, Unger R, Auerbach D, Given JA, Moulton J (1997) Chaos in protein dynamics. *PROTEINS: Struct Funct Genet* 29:417–425
32. Zhou H, Wang L (1996) Chaos in biomolecular dynamics. *J Phys Chem* 100:8101–8105
33. Duan Y, Kollman PA (2001) Computational protein folding: from lattice to all-atom. *IBM Syst J* 40:297–309